Microbiol. Biotechnol. Lett. (2023), 51(4), 538–541
http://dx.doi.org/10.48022/mbl.2311.11003
pISSN 1598-642X eISSN 2234-7305

MBL
Microbiology and Biotechnology Letters

Genome Reports

Draft Genome Sequence of *Aureobasidium pullulans* Strain MHAU2101, a Biological Control Agent against Fire Blight from Korea

Lin He[1], Huan Luo[1], Mi-Hyun Lee[2], and Jun Myoung Yu[1]*

[1]*Department of Applied Biology, Chungnam National University, Daejeon 34134, Republic of Korea*
[2]*Crop Protection Division, National Institute of Agricultural Sciences, Wanju 55365, Republic of Korea*

Received: November 7, 2023 / Revised: November 26, 2023/ Accepted: November 28, 2023

In this study, we present the draft genome of *Aureobasidium pullulans* strain MHAU2101, which is the first strain to effectively control fire blight caused by *Erwinia amylovora* in Korea. The genome of strain MHAU2101 was composed of 28,669,322 base pairs, with a C+G content of 50.4%. The assembly comprised 17 contigs and had 99.22% completeness. The results of this study will be a valuable resource for future research on the biocontrol mechanism of *A. pullulans* strain MHAU2101.

Keywords: *Aureobasidium pullulans*, biocontrol, fire blight, PacBio, Illumina, draft genome

Aureobasidium pullulans is often referred to as "yeast-like" or "black yeast" due to its melanized yeast-like morphology. This fungus is ubiquitous in a variety of environments and is frequently isolated from the phyllosphere and soil [1]. *A. pullulans* is well-known for its significant biotechnological potential in medicine, pharmacy, and the food industry, which has been attributed to its capacity to produce polymeric substances, particularly pullulan, and hydrolytic enzymes [2, 3]. Additionally, this fungus has been used as a commercial biocontrol agent against postharvest diseases and fire blight [4, 5]. Notably, strain MHAU2101 was recently isolated from pears in Korea and demonstrated great efficacy in controlling fire blight caused by *Erwinia amylovora* [6]. The fungal strain MHAU2101 was identified as *A. pullulans* by morphological and phylogenetic analyses, employing large-subunit 26S ribosomal DNA and the internal transcribed spacer [6].

In this study, we sequenced the genome of *A. pullulans* MHAU2101, isolated from pear flowers in Wanju, South Korea, in 2021 [6]. The genomic DNA of strain MHAU2101 was extracted using the Quick-DNA Fungal/Bacterial Miniprep Kit (Zymo Research, USA). DNA purity and content were evaluated by spectrophotometry (NanoPhotometer NP80, Implen, Germany), and DNA integrity was assessed with electrophoresis. The quality-controlled DNA was used for genome sequencing using a combination of the PacBio Sequel and Illumina NovaSeq sequencing platforms, and the PacBio SMRTbell prep kit 3.0 (PacBio, USA) and TruSeq Nano DNA Preparation Kit (350) (Illumina, USA) were used to construct two sequencing libraries, respectively. Sequencing was performed by SEEDERS Co. (Republic of Korea; http://www.seeders.co.kr/). The Illumina data were used to polish the assembled genome using PacBio data and Pilon v1.21 software [7]. The benchmarking universal single-copy orthologs (BUSCO v5.1.3) dataset eukaryota_

***Corresponding author**
Phone: +82-42-821-5768
E-mail: junyu@cnu.ac.kr

odb10 was used to evolve the completeness of the assembled genome [8]. The MAKER v3.01.03 program [9] was used to predict the protein-coding genes. Additionally, a gene functional analysis was performed using the Gene Ontology (GO), InterPro (v69.0), Pfam (v31.0), Conserved Domains Database (CDD v3.16), TIGRFAM (v15.0), and EggNOG (v4.5.1) databases. In this study, default parameters were utilized for all software and databases.

In total, 140,079 HiFi reads were obtained from PacBio sequencing, with a HiFi N_{50} value of 8,264 bp (half of all bases resided in reads of this size or longer), and an average read length of 7,301 bp, consisting of 1,022,750,246 HiFi bases. Illumina sequencing generated approximately 2,022.9 Mb of raw data, and 1,364.2 Mb of clean data were retained after filtering. *De novo* assembly was performed on the pure PacBio long reads using SMART Link v11.1.0 (https://www.pacb.com/), and an automatic circular check was conducted to ensure there were no overlaps within contigs. Illumina short reads were applied for sequence compensation to construct the contigs more accurately. Read mapping and data polishing were performed using Racon v1.4.13-74c937c program [10]. The classification of the species has been confirmed as *A. pullulans* MHAU2101 based on the whole-genome data. Upon comparing this data with the strain *A. pullulans* EXF-150 (Accession: AYEO00000000.1), it was observed that they share an average nucleotide identity (ANI) of 97.28% and a Digital DNA-DNA hybridization (dDDH) coverage of 73.70%.

The assembled draft genome was 28,669,322 bp in size, comprising 17 contigs and 50.4% G+C content (Fig. 1). The N_{50} of the genome was 2,213,608 bp. The length of the 17 contigs ranged from 18,956 bp to 4,494,408 bp, and the average length of a contig was 1,686,430 bp. The BUSCO results indicated that the genome assembly displayed a high overall completeness of 99.22%; only 0.78% of BUSCOs were present as fragments, and none were missing in the analysis. The

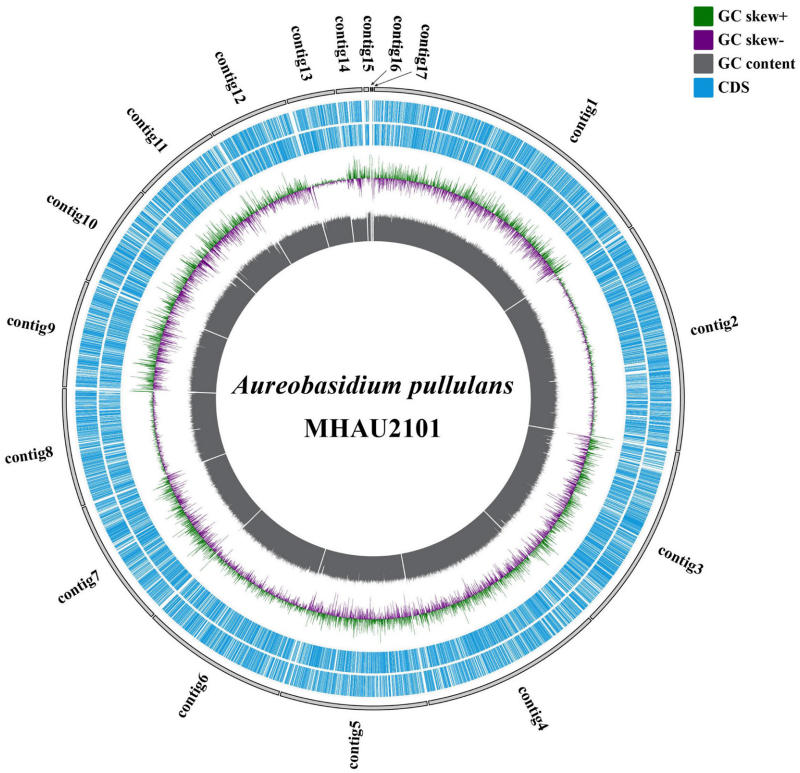

Fig. 1. Circular representation of the draft genome of *Aureobasidium pullulans* MHAU2101. From the outside to the inside of the circle graph are the genome contigs, the predicted coding sequence (sense strand and antisense strand), GC skew, positive (green), negative (purple), and the inner circle indicates the GC content. All these contigs are visualized in a single circular representation. In the assembly, contig 16 is the only one that is circular.

Table 1. Genome statistics of *Aureobasidium pullulans* Strain MHAU2101.

Features	Strain MHAU2101
PacBio platform	
Number of read	140,079
Total length of reads	1,022,750,246 bp
N50 length	8,264 bp
Average read length	7,301 bp
Illumina platform	
Raw data of read	13,396,436
Clean data of read	9,037,346
Clean data Q20	99.32%
Clean data Q30	96.78%
Clean data GC	49.59%
Genome assembly	
Contig	17
Contig size	28,669,322 bp
N50 length	2,213,608 bp
GC content	50.4%
Number of exons	32, 338
Number of introns	20,765
Putative protein-coding genes	11,573
Number of tRNA	313
Number of rRNA	70
BUSCOs	
Complete and single-copy BUSCOs	253 (99.22%)
Complete and duplicated BUSCOs	0 (0.00%)
Fragmented BUSCOs	2 (0.78%)
Missing BUSCOs	0 (0.00%)
Total BUSCO groups searched	255 (100%)

genome contained 20,765 introns and 32,338 exons in total. Moreover, 11,573 protein-coding genes, 313 transfer RNAs, and 70 ribosomal RNA coding sequences were predicted. The features of the strain MHAU2101 genome are shown in Table 1. A total of 6,043 proteins were assigned to the GO database, 7,664 proteins were assigned to the InterPro database, 7,864 proteins were assigned to the Pfam database, 3,511 proteins were assigned to the CDD database, 878 proteins were assigned to the TIGRFAM database, and 11,054 proteins were assigned to the EggNOG database.

In summary, we present a high-quality draft genome assembly of *A. pullulans* MHAU2101, a biocontrol strain known for its efficacy against fire blight caused by *E. amylovora*. The insights derived from this genome assembly not only enhance our fundamental understanding but also carry practical implications for developing targeted strategies to combat the devastating plant disease, fire blight. The genomic sequence serves as a pivotal resource for future research on genetic diversity and biocontrol mechanisms of *A. pullulans*. This will significantly enrich the knowledge base essential for advancing our understanding of biocontrol mechanisms and facilitating the development of practical applications in agriculture.

Nucleotide Sequence Accession Number(s)

The draft genome sequence of *A. pullulans* strain MHAU2101 in this article was deposited at DDBJ/ENA/GenBank under accession number JAWJBJ000000000. The version described in this study is JAWJBJ010000000. (BioProject: PRJNA1027605; BioSample: SAMN37801316).

Acknowledgments

This work was supported by a grant from the Agenda program (PJ015594) of the Rural Development Administration, Republic of Korea.

Conflict of Interest

The authors have no financial conflicts of interest to declare.

References

1. Urzì C, De Leo F, Passo CL, Criseo G. 1999. Intra-specific diversity of *Aureobasidium pullulans* strains isolated from rocks and other habitats assessed by physiological methods and by random amplified polymorphic DNA (RAPD). *J. Microbiol. Methods* **36**: 95-105.
2. Chi Z, Wang F, Chi Z, Yue L, Liu G, Zhang T. 2009. Bioproducts from *Aureobasidium pullulans*, a biotechnologically important yeast. *Appl. Microbiol. Biotechnol.* **82**: 793-804.
3. Ravella SR, Quiñones TS, Retter A, Heiermann M, Amon T, Hobbs PJ. 2010. Extracellular polysaccharide (EPS) production by a novel strain of yeast-like fungus *Aureobasidium pullulans*. *Carbohydr. Polym.* **82**: 728-732.
4. Zhang D, Spadaro D, Garibaldi A, Gullino ML. 2010. Efficacy of the antagonist *Aureobasidium pullulans* PL5 against postharvest pathogens of peach, apple and plum and its modes of action. *Biol. Control* **54**: 172-180.
5. Kunz S, Zeng Q, Johnson KB. 2023. History, efficacy, orchard ecology, and mode of action of *Aureobasidium pullulans*, the microbial agent in Blossom Protect, for suppression of fire blight of pome fruit. *J. Plant. Pathol.* https://doi.org/10.1007/s42161-023-

01448-4.
6. Oh H, Choi H-W, Lee YH, Oh S-K. 2022. Inhibitory effects of *Aureobasidium pullulans* MHAU2101 isolated from domestic pear blossom against fire blight. *Korean J. Mycol.* **50**: 332.
7. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, *et al.* 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* **9**: e112963.
8. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**: 3210-3212.
9. Brůna T, Hoff KJ, Lomsadze A, Stanke M, Borodovsky M. 2021. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR. Genom. Bioinform.* **3**: lqaa108.
10. Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* **27**: 737-746.